19. Explain about k-means clustering. Give its significance.

20. Describe about how association rules are used in data analytics in detail.

———————

APRIL/MAY 2025

23PCS34 — DATA SCIENCE AND ANALYTICS

Time : Three hours Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

Shanmuga Industries Arts & Science College T.V.Malai

1. What is Data science?

2. Define the term - machine learning.

3. What is Data analytics?

4. List the various technologies used in data analytics.

5. Give an overview of R Studio.

6. Categorize the types of attributes in R.

7. Differentiate clustering and classification.

8. Define the Bayes' theorem.

9. What is Deep learning?

10. How Artificial intelligence is used in Data science?

Shanmuga Industries Arts & Science College T.V.Malai

SECTION B — (5 × 5 = 25 marks)

Answer ALL questions.

11. (a) List the facets of Data in Data science and explain any two.

Or

(b) Explain the different types of Machine learning with suitable examples.

12. (a) Discuss on Data analytics. Give the significance of Data analytics life cycle.

Or

(b) Write about the advanced Data analytics tools.

13. (a) What is dirty data in data analytics? Explain.

Or

(b) How to get descriptive statistics in R for multiple variables? Discuss.

14. (a) How evaluation is performed on Decision trees? Explain.

Or

(b) Illustrate Clustering analysis methods.

15. (a) Discuss on Fuzzy logic.

Or

(b) Explain Linear regression with example.

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Describe the concepts of Data science process.

17. Explain life cycle phases of Data analytics.

18. How to perform Exploratory Data analytics? Give example.